



All Databases

PubMed

Nucleotide

Protein

Genome

Structure

OMIM

PMC

Journals

Books

Search PubMed

for flux[au]

Preview

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

About Entrez

Text Version

Entrez PubMed

Overview

Help | FAQ

Tutorial

New/Noteworthy

E-Utilities

PubMed Services

Journals Database

MeSH Database

Single Citation

Matcher

Batch Citation Matcher

Clinical Queries

Special Queries

LinkOut

My NCBI (Cubby)

Related Resources

Order Documents

NLM Catalog

NLM Gateway

TOXNET

Consumer Health

Clinical Alerts

ClinicalTrials.gov

PubMed Central

- Search History will be lost after eight hours of inactivity.
- To combine searches use # before search number, e.g., #2 AND #6.
- Search numbers may not be continuous; all searches are represented.
- Click on query # to add to strategy

| Search | Most Recent Queries | Time | Result |
|--------|---|----------|---------|
| #24 | Search flux[au] | 14:54:48 | 1 |
| #23 | Search flux m[au] | 14:54:36 | 0 |
| #22 | Search drittani l[au] | 14:54:02 | 12 |
| #21 | Search drittani l[au] | 14:53:55 | 0 |
| #20 | Search #19 and #3 | 14:53:34 | 2 |
| #19 | Search vega m[au] | 14:53:17 | 324 |
| #5 | Search #4 and #3 Sort by: PublicationDate | 14:42:35 | 55 |
| #9 | Search #3 and #7 Sort by: PublicationDate | 14:37:33 | 31 |
| #8 | Search #4 and #7 | 14:37:14 | 0 |
| #7 | Search (increased or enhanced or augmented or elevated) and titer | 14:36:57 | 4596 |
| #6 | Search increased or enhanced or augmented or elevated | 14:36:42 | 1511334 |
| #4 | Search mutant rep | 14:36:03 | 192 |
| #3 | Search adeno-associated virus | 14:35:56 | 2330 |
| #2 | Search adeno associated virus | 14:35:50 | 2330 |
| #1 | Search adenoassociated virus | 14:35:45 | 61 |

Clear History

Write to the Help Desk

NCBI | NLM | NIH

Department of Health & Human Services

Privacy Statement | Freedom of Information Act | Disclaimer

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2005, 04:41:56 ; Search time 175 Seconds
(without alignments)
1372.447 Million cell updates/sec

Title: US-10-022-390-113
Perfect score: 3312
Sequence: 1 TAGFYEIVIKVPSDLDEHLP.....DACTACDLVNVLDLDCIFEQ 621

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 3312 | 100.0 | 621 | 6 | ABU97503 Adeno ass |
| 2 | 3312 | 100.0 | 621 | 7 | ADB70532 Adeno-ass |
| 3 | 3304 | 99.8 | 621 | 6 | ABU97507 Adeno ass |
| 4 | 3304 | 99.8 | 621 | 7 | ADB70536 Adeno-ass |
| 5 | 3303 | 99.7 | 621 | 6 | ABU97559 Adeno ass |
| 6 | 3303 | 99.7 | 621 | 6 | ABU97643 Adeno ass |
| 7 | 3303 | 99.7 | 621 | 6 | ABU97671 Adeno ass |
| 8 | 3303 | 99.7 | 621 | 6 | ABU97547 Adeno ass |
| 9 | 3303 | 99.7 | 621 | 6 | ABU97647 Adeno ass |
| 10 | 3303 | 99.7 | 621 | 6 | ABU97659 Adeno ass |
| 11 | 3303 | 99.7 | 621 | 6 | ABU97439 Adeno ass |
| 12 | 3303 | 99.7 | 621 | 6 | ABU97651 Adeno ass |
| 13 | 3303 | 99.7 | 621 | 6 | ABU97441 Adeno ass |
| 14 | 3303 | 99.7 | 621 | 6 | ABU97543 Adeno ass |
| 15 | 3303 | 99.7 | 621 | 6 | ABU97635 Adeno ass |
| 16 | 3303 | 99.7 | 621 | 7 | ADB70680 Adeno-ass |
| 17 | 3303 | 99.7 | 621 | 7 | ADB70688 Adeno-ass |
| 18 | 3303 | 99.7 | 621 | 7 | ADB70576 Adeno-ass |
| 19 | 3303 | 99.7 | 621 | 7 | ADB70468 Adeno-ass |
| 20 | 3303 | 99.7 | 621 | 7 | ADB70700 Adeno-ass |
| 21 | 3303 | 99.7 | 621 | 7 | ADB70572 Adeno-ass |
| 22 | 3303 | 99.7 | 621 | 7 | ADB70672 Adeno-ass |
| 23 | 3303 | 99.7 | 621 | 7 | ADB70588 Adeno-ass |
| 24 | 3303 | 99.7 | 621 | 7 | ADB70664 Adeno-ass |
| 25 | 3303 | 99.7 | 621 | 7 | ADB70676 Adeno-ass |

| | | | | | | |
|----|------|------|-----|---|----------|--------------------|
| 26 | 3303 | 99.7 | 621 | 7 | ADB70470 | Adb70470 Adeno-ass |
| 27 | 3302 | 99.7 | 621 | 6 | ABU97527 | Abu97527 Adeno ass |
| 28 | 3302 | 99.7 | 621 | 6 | ABU97443 | Abu97443 Adeno ass |
| 29 | 3302 | 99.7 | 621 | 6 | ABU97685 | Abu97685 Adeno ass |
| 30 | 3302 | 99.7 | 621 | 6 | ABU97655 | Abu97655 Adeno ass |
| 31 | 3302 | 99.7 | 621 | 6 | ABU97437 | Abu97437 Adeno ass |
| 32 | 3302 | 99.7 | 621 | 7 | ADB70466 | Adb70466 Adeno-ass |
| 33 | 3302 | 99.7 | 621 | 7 | ADB70472 | Adb70472 Adeno-ass |
| 34 | 3302 | 99.7 | 621 | 7 | ADB70684 | Adb70684 Adeno-ass |
| 35 | 3302 | 99.7 | 621 | 7 | ADB70556 | Adb70556 Adeno-ass |
| 36 | 3302 | 99.7 | 621 | 7 | ADB70714 | Adb70714 Adeno-ass |
| 37 | 3301 | 99.7 | 621 | 6 | ABU97491 | Abu97491 Adeno ass |
| 38 | 3301 | 99.7 | 621 | 6 | ABU97719 | Abu97719 Adeno ass |
| 39 | 3301 | 99.7 | 621 | 6 | ABU97397 | Abu97397 Adeno ass |
| 40 | 3301 | 99.7 | 621 | 6 | ABU97447 | Abu97447 Adeno ass |
| 41 | 3301 | 99.7 | 621 | 6 | ABU97405 | Abu97405 Adeno ass |
| 42 | 3301 | 99.7 | 621 | 6 | ABU97551 | Abu97551 Adeno ass |
| 43 | 3301 | 99.7 | 621 | 6 | ABU97599 | Abu97599 Adeno ass |
| 44 | 3301 | 99.7 | 621 | 6 | ABU97499 | Abu97499 Adeno ass |
| 45 | 3301 | 99.7 | 621 | 6 | ABU97567 | Abu97567 Adeno ass |

ALIGNMENTS

RESULT 1
ABU97503
ID ABU97503 standard; protein; 621 AA.
XX
AC ABU97503;
XX
XX
DT 30-JUL-2003 (first entry)
XX
DE Adeno associated virus (AAV) mutant rep protein #113.
XX
KW Polypeptide production; protein production; target protein;
KW high throughput directed evolution; rational mutagenesis; AAV;
KW protein variant generation; virus titering; Adeno associated virus;
KW rep protein; mutant; mutein.
XX
OS Adeno associated virus.
OS Synthetic.
XX
PN WO2003023032-A2.
XX
PD 20-MAR-2003.
XX
PF 16-AUG-2002; 2002WO-IB003921.
PR 27-AUG-2001; 2001US-0315382P.
PR 17-DEC-2001; 2001US-00022249.
XX
PA (NAUT-) NAUTILUS BIOTECH.
XX
PI Vega M, Drittanti L, Flaux M;
XX
DR WPI; 2003-354538/33.
XX
PT Producing proteins with a predetermined property comprises introducing
PT nucleic acids encoding a modified target protein into host cells and
PT expressing and screening for proteins with different activity than the
PT target protein.
XX
PS Disclosure; SEQ ID NO 113; 141pp; English.
XX
CC The present invention relates to method for producing peptides,
CC polypeptides, or proteins having a predetermined property. The method
CC comprises (a) producing a population of sets of polynucleotide sequences
CC that encode modified forms of a target protein, (b) introducing each set
CC of polynucleotide sequences into host cells and expressing the encoded
CC protein, where the host cells are present in an addressable array, and
CC (c) individually screening the sets of encoded proteins. The method is
CC useful in performing high throughput directed evolution of peptides and

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2005, 04:42:27 ; Search time 43 Seconds
(without alignments)
1389.549 Million cell updates/sec

Title: US-10-022-390-113
Perfect score: 3312
Sequence: 1 TAGFYEIVIKVPSDLDEHLP.....DACTACDLVNVLDLDCIFEQ 621

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 2771 | 83.7 | 536 | 1 UYAD1A | noncapsid protein |
| 2 | 1470 | 44.4 | 626 | 2 S52209 | noncapsid protein |
| 3 | 602 | 18.2 | 671 | 1 UYPV19 | noncapsid protein |
| 4 | 482.5 | 14.6 | 490 | 2 T44050 | hypothetical prote |
| 5 | 479.5 | 14.5 | 490 | 2 JQ1630 | noncapsid protein |
| 6 | 457 | 13.8 | 660 | 1 UYPVPP | noncapsid protein |
| 7 | 456 | 13.8 | 662 | 1 UYPVNA | noncapsid protein |
| 8 | 446 | 13.5 | 726 | 1 UYPVS1 | noncapsid protein |
| 9 | 443.5 | 13.4 | 668 | 1 UYPVCP | noncapsid protein |
| 10 | 441.5 | 13.3 | 668 | 1 UYPVME | noncapsid protein |
| 11 | 439.5 | 13.3 | 668 | 1 UYPVFP | noncapsid protein |
| 12 | 433.5 | 13.1 | 668 | 1 A44276 | noncapsid protein |
| 13 | 429 | 13.0 | 392 | 1 UYPV1F | noncapsid protein |
| 14 | 427.5 | 12.9 | 672 | 1 UYPV1 | noncapsid protein |
| 15 | 426.5 | 12.9 | 721 | 1 UYPVIM | noncapsid protein |
| 16 | 421.5 | 12.7 | 672 | 1 UYPVIM | noncapsid protein |
| 17 | 347.5 | 10.5 | 276 | 2 S26428 | hypothetical 31.5K |
| 18 | 330.5 | 10.0 | 641 | 2 S41861 | gene NS-1 protein |
| 19 | 325.5 | 9.8 | 641 | 2 S41439 | gene NS-1 protein |
| 20 | 315.5 | 9.5 | 641 | 2 S41434 | gene NS-1 protein |
| 21 | 303 | 9.1 | 620 | 1 UYPVAP | noncapsid protein |
| 22 | 148 | 4.5 | 849 | 1 UYPVAD | noncapsid protein |
| 23 | 142 | 4.3 | 885 | 1 VCPVF2 | structural protein |
| 24 | 141.5 | 4.3 | 545 | 2 B44054 | orf2 protein - Jun |
| 25 | 139.5 | 4.2 | 646 | 2 S36586 | E1 protein - human |
| 26 | 139 | 4.2 | 647 | 2 S36575 | E1 protein - human |
| 27 | 136.5 | 4.1 | 625 | 1 W1WLR1 | E1 protein - rhesu |
| 28 | 135 | 4.1 | 644 | 1 W1WL33 | E1 protein - human |
| 29 | 135 | 4.1 | 644 | 1 W1WL58 | E1 protein - human |

| | | | | | |
|----|-------|-----|------|----------|--------------------|
| 30 | 132 | 4.0 | 611 | 1 W1WLEP | E1 protein - Europ |
| 31 | 131.5 | 4.0 | 631 | 2 S36505 | E1 protein - human |
| 32 | 129 | 3.9 | 647 | 2 S36517 | E1 protein - human |
| 33 | 127.5 | 3.8 | 620 | 1 W1WLB2 | E1 protein - bovin |
| 34 | 126.5 | 3.8 | 638 | 2 S36546 | E1 protein - human |
| 35 | 125 | 3.8 | 605 | 1 W1WLEB | E1 protein - bovin |
| 36 | 124.5 | 3.8 | 643 | 2 S36563 | E1 protein - human |
| 37 | 124.5 | 3.8 | 869 | 2 H86440 | unknown protein [i |
| 38 | 119 | 3.6 | 881 | 1 B71316 | endopeptidase La (|
| 39 | 117.5 | 3.5 | 802 | 2 T23295 | hypothetical prote |
| 40 | 117.5 | 3.5 | 860 | 2 T23296 | hypothetical prote |
| 41 | 117 | 3.5 | 648 | 1 W1WLC1 | E1 protein - pygmy |
| 42 | 116 | 3.5 | 647 | 2 S36557 | E1 protein - human |
| 43 | 116 | 3.5 | 681 | 2 S36534 | E1 protein - human |
| 44 | 115 | 3.5 | 647 | 1 W1WL39 | E1 protein - human |
| 45 | 114.5 | 3.5 | 4725 | 1 A44357 | dynein heavy chain |

ALIGNMENTS

RESULT 1

UYAD1A
noncapsid protein NS1 - adeno-associated virus type 2
N;Contains: noncapsid protein NS2
C;Species: adeno-associated virus type 2
C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004
C;Accession: A03694
R;Srivastava, A.; Lusby, E.W.; Berns, K.I.
J. Virol. 45, 555-564, 1983
A;Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.
A;Reference number: A03694; MUID:83164299; PMID:6300419
A;Accession: A03694
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-536 <SRI>
A;Cross-references: UNIPROT:P03132; EMBL:J01901; NID:G209616; PIDN:AAA42372.1; PID:G209
C;Genetics:
A;Introns: 529/2
C;Superfamily: parvovirus noncapsid protein
C;Keywords: noncapsid protein
F;225-536/Product: noncapsid protein NS2 #status predicted <NS2>

Query Match 83.7%; Score 2771; DB 1; Length 536;
Best Local Similarity 99.6%; Pred. No. 2.3e-186;
Matches 525; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 3 | GFYEIVIKVPSDLDEHLPGISDSFVNWVAEKWEWELPPDSDMDNLIEQAPLTVAEKLQRD | 62 |
| Db | 3 | GFYEIVIKVPSDLDGHLPGISDSFVNWVAEKWEWELPPDSDMDNLIEQAPLTVAEKLQRD | 62 |
| QY | 63 | FLTEWRRVSKAPEALFFVQFEKGESYFHMVHLVETTGKSMVLGRFLSQIREKLIQRIYR | 122 |
| Db | 63 | FLTEWRRVSKAPEALFFVQFEKGESYFHMVHLVETTGKSMVLGRFLSQIREKLIQRIYR | 122 |
| QY | 123 | GIEPTLPNWFVAVTKTRNGAGGKVKVDECYIPNYLLPKTQPELQAWTNMEOYLSACLNL | 182 |
| Db | 123 | GIEPTLPNWFVAVTKTRNGAGGKVKVDECYIPNYLLPKTQPELQAWTNMEOYLSACLNL | 182 |
| QY | 183 | TERKRLVAQHLTHVSQTQEQNKENQNPNSDAPVIRSKTSARYMELVGLVDKGITSEKQW | 242 |
| Db | 183 | TERKRLVAQHLTHVSQTQEQNKENQNPNSDAPVIRSKTSARYMELVGLVDKGITSEKQW | 242 |
| QY | 243 | IQEDQASYISFNAASNSRSQIKAAALDNAGKIMSLTKTAPDYLVGQQPVEDISSNRIYKIL | 302 |
| Db | 243 | IQEDQASYISFNAASNSRSQIKAAALDNAGKIMSLTKTAPDYLVGQQPVEDISSNRIYKIL | 302 |
| QY | 303 | ELNGYDPPQYAAASVFLGWATKKFGKRNTIWLFGPATTKGTNTIAEIAHNVPFYGCNVWTNE | 362 |
| Db | 303 | ELNGYDPPQYAAASVFLGWATKKFGKRNTIWLFGPATTKGTNTIAEIAHNVPFYGCNVWTNE | 362 |
| QY | 363 | NFPFNDCCVDKMWIWEEGKMTAKVVESAKAILGGSKVRVDQCKSSAQIDPTPVI VTSNT | 422 |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2005, 04:42:07 ; Search time 177 Seconds
(without alignments)
1796.617 Million cell updates/sec

Title: US-10-022-390-113
 Perfect score: 3312
 Sequence: 1 TAGFEIVIKVPSDLDEHLP.....DACTACDLVNVDLDDCIFEQ 621

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 16123378

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10
                  Listing first 45

```

```
Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|-----------|--------------------|
| 1 | 3297 | 99.5 | 621 | 2 | O56651 | O56651 adeno-assoc |
| 2 | 3290 | 99.3 | 621 | 2 | Q89268 | Q89268 adeno-assoc |
| 3 | 3277 | 98.9 | 621 | 2 | Q670R3 | Q670r3 adeno-assoc |
| 4 | 3276 | 98.9 | 621 | 2 | Q670Q9 | Q670q9 adeno-assoc |
| 5 | 3275 | 98.9 | 621 | 2 | Q670R7 | Q670r7 adeno-assoc |
| 6 | 3268 | 98.7 | 621 | 2 | Q670R1 | Q670r1 adeno-assoc |
| 7 | 3263 | 98.5 | 621 | 2 | Q670R5 | Q670r5 adeno-assoc |
| 8 | 3034 | 91.6 | 623 | 2 | O41854 | O41854 adeno-assoc |
| 9 | 2992.5 | 90.4 | 624 | 2 | O56138 | O56138 adeno-assoc |
| 10 | 2974.5 | 89.8 | 624 | 2 | Q65310 | Q65310 adeno-assoc |
| 11 | 2974 | 89.8 | 623 | 2 | Q8JQG1 | Q8jqg1 adeno-assoc |
| 12 | 2954 | 89.2 | 623 | 2 | Q9WBP7 | Q9wbp7 adeno-assoc |
| 13 | 2949 | 89.0 | 623 | 2 | O56136 | O56136 adeno-assoc |
| 14 | 2851 | 86.1 | 625 | 2 | Q8JQF9 | Q8jqf9 adeno-assoc |
| 15 | 2778 | 83.9 | 536 | 2 | O56650 | O56650 adeno-assoc |
| 16 | 2771 | 83.7 | 536 | 1 | VNCA_AAV2 | P03132 adeno-assoc |
| 17 | 2122 | 64.1 | 397 | 2 | Q89270 | Q89270 adeno-assoc |
| 18 | 1913.5 | 57.8 | 610 | 2 | Q6JL80 | Q6jl80 bovine aden |
| 19 | 1906.5 | 57.6 | 610 | 2 | Q9YJC1 | Q9yjc1 adeno-assoc |
| 20 | 1876 | 56.6 | 663 | 2 | Q6GWF3 | Q6gwf3 avian adeno |
| 21 | 1824.5 | 55.1 | 662 | 2 | Q7TG44 | Q7tg44 avian adeno |
| 22 | 1603 | 48.4 | 312 | 2 | Q89269 | Q89269 adeno-assoc |
| 23 | 1486.5 | 44.9 | 627 | 2 | Q65443 | Q65443 muscovy duc |
| 24 | 1481 | 44.7 | 627 | 2 | Q67665 | Q67665 goose parvo |
| 25 | 1478 | 44.6 | 627 | 2 | Q6R968 | Q6r968 goose parvo |
| 26 | 1474 | 44.5 | 627 | 2 | Q8V396 | Q8v396 goose parvo |
| 27 | 1470 | 44.4 | 626 | 2 | Q83288 | Q83288 muscovy duc |
| 28 | 1138.5 | 34.4 | 461 | 2 | Q67671 | Q67671 goose parvo |
| 29 | 943 | 28.5 | 562 | 2 | Q6V7U3 | Q6v7u3 snake parvo |
| 30 | 716.5 | 21.6 | 711 | 2 | P87583 | P87583 chipmunk pa |
| 31 | 696 | 21.0 | 537 | 2 | Q918V2 | Q918v2 bovine parv |

RESULT 1

| | | | |
|--------|--|----------------------------------|---------|
| OS6651 | PRELIMINARY; | PRT; | 621 AA. |
| ID | O56651 | | |
| AC | O56651; | | |
| DT | 01-JUN-1998 | (TREMBLrel. 06, Created) | |
| DT | 01-JUN-1998 | (TREMBLrel. 06, Last sequence up | |
| DT | 01-MAR-2004 | (TREMBLrel. 26, Last annotation | |
| DE | Rep 78 protein. | | |
| OS | Adeno-associated virus 2 (AAV2). | | |
| OC | Viruses; ssDNA viruses; Parvoviridae; Parvov | | |
| OX | NCBI TaxID=10804; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RX | MEDLINE=95088582; PubMed=7996133; | | |
| RA | Ruffing M., Heid H., Kleinschmidt J.A.; | | |
| RT | "Mutations in the carboxy terminus of adeno- | | |
| RT | proteins affect viral infectivity: lack of a | | |
| RT | motif."; | | |
| RL | J. Gen. Virol. 75:0-0(0). | | |
| RP | [2] | | |
| RP | SEQUENCE FROM N.A. | | |
| RA | Berns K.I., Bohenzky R.A., Cassinotti P., Co | | |
| RA | Dull T., Horer M., Kleinschmidt J.A., Ruffin | | |
| RA | Tratschin J.-D., Weitz M.; | | |
| RL | Submitted (JAN-1998) to the EMBL/GenBank/DBS | | |
| DR | EMBL; AF043303; AAC03775.1; .- | | |
| DR | HSSP; Q9YJC1; 1M55. | | |
| DR | GO; GO:0019079; P:viral genome replication; | | |
| DR | InterPro; IPR001257; Parvo NS1. | | |
| DR | Pfam; PF01057; Parvo NS1; 1. | | |
| SQ | SEQUENCE 621 AA; 70635 MW; 14BF05D4AED1A | | |

Query Match 99.5%; Score 3297; DB 2; Length 621;
Best Local Similarity 99.8%; Pred. No. 6.2e-220;
Matches 618; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

| | | | |
|----|-----|--|------|
| Qy | 3 | GFYEIVIKVPSDLDEHLPGISDSFVNWVAEKEWELPPDSDMDLNLIEQAPLTVAEKQLRD | 62 |
| Dd | 3 | GFYEIVIKVPSDLDEHLPGISDSFVNWVAEKEWELPPDSDMDLNLIEQAPLTVAEKQLRD | 62 |
| Qy | 63 | FLTEWRRVSKAPEALFFVQPEKGESYFHMHVLVETTGKSMVLGRFLSQIREKLIQRIYR | 1222 |
| Dd | 63 | FLTEWRRVSKAPEALFFVQPEKGESYFHMHVLVETTGKSMVLGRFLSQIREKLIQRIYR | 1222 |
| Qy | 123 | GIEPTLPNWFAVTKTRNGAGGKNKVDECIIPNYLLPKTQPELQAWTNNMEQYLSACLNL | 1822 |
| Dd | 123 | GIEPTLPNWFAVTKTRNGAGGKNKVDECIIPNYLLPKTQPELQAWTNNMEQYLSACLNL | 1822 |
| Qy | 183 | TERKRLVAQHLTHVSTQEQNKENQNPNSDAPVIRSKTSARYMELVGLVDKGITSEKQW | 2422 |
| Dd | 183 | TERKRLVAQHLTHVSTQEQNKENQNPNSDAPVIRSKTSARYMELVGLVDKGITSEKQW | 2422 |
| Qy | 243 | IQEDQASYISFNAASNRSQIKAALDNAGKIMSLTKTAPDYLVGQPVEDISSNRIYKIL | 302 |

ALIGNMENTS

| | | | | | | | |
|----|-------|------|-----|---|--------|--------|-------------|
| 32 | 671.5 | 20.3 | 672 | 2 | Q9J0X5 | Q9j0x5 | pig-tailed |
| 33 | 645.5 | 19.5 | 687 | 2 | Q88271 | Q88271 | simian parv |
| 34 | 637 | 19.2 | 661 | 2 | Q8QVL2 | Q8qvl2 | hamster par |
| 35 | 632 | 19.1 | 651 | 2 | Q918V0 | Q918v0 | bovine parv |
| 36 | 631.5 | 19.1 | 683 | 2 | Q9J0X7 | Q9j0x7 | rhesus maca |
| 37 | 614 | 18.5 | 671 | 2 | Q9WKM1 | Q9wkm1 | human parvo |
| 38 | 612 | 18.5 | 671 | 2 | Q75U94 | Q75u94 | human parvo |
| 39 | 611 | 18.4 | 658 | 2 | Q9PZT5 | Q9pzt5 | human parvo |
| 40 | 611 | 18.4 | 671 | 2 | Q85116 | Q85116 | human parvo |
| 41 | 609 | 18.4 | 671 | 2 | Q9PZT1 | Q9pzt1 | human parvo |
| 42 | 608 | 18.4 | 662 | 2 | Q9PZU1 | Q9pzu1 | human parvo |
| 43 | 608 | 18.4 | 671 | 2 | Q75U76 | Q75u76 | human parvo |
| 44 | 607 | 18.3 | 671 | 2 | Q75U85 | Q75u85 | human parvo |
| 45 | 606 | 18.3 | 659 | 2 | Q9PZT7 | Q9pzt7 | human parvo |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2005, 04:59:05 ; Search time 135 Seconds
(without alignments)
1505.309 Million cell updates/sec

Title: US-10-022-390-113
Perfect score: 3312
Sequence: 1 TAGFYEIVIKVPSDLDEHLP.....DACTACDLNVNVDLDDCIFEQ 621

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 3312 | 100.0 | 621 | 14 | US-10-022-390-113 |
| 2 | 3312 | 100.0 | 621 | 14 | US-10-022-249-113 |
| 3 | 3304 | 99.8 | 621 | 14 | US-10-022-390-117 |
| 4 | 3304 | 99.8 | 621 | 14 | US-10-022-249-117 |
| 5 | 3303 | 99.7 | 621 | 14 | US-10-022-390-49 |
| 6 | 3303 | 99.7 | 621 | 14 | US-10-022-390-51 |
| 7 | 3303 | 99.7 | 621 | 14 | US-10-022-390-153 |
| 8 | 3303 | 99.7 | 621 | 14 | US-10-022-390-157 |
| 9 | 3303 | 99.7 | 621 | 14 | US-10-022-390-169 |
| 10 | 3303 | 99.7 | 621 | 14 | US-10-022-390-245 |
| 11 | 3303 | 99.7 | 621 | 14 | US-10-022-390-253 |
| 12 | 3303 | 99.7 | 621 | 14 | US-10-022-390-257 |
| 13 | 3303 | 99.7 | 621 | 14 | US-10-022-390-261 |

| | | | | | | |
|----|------|------|-----|----|-------------------|-------------------|
| 14 | 3303 | 99.7 | 621 | 14 | US-10-022-390-269 | Sequence 269, App |
| 15 | 3303 | 99.7 | 621 | 14 | US-10-022-390-281 | Sequence 281, App |
| 16 | 3303 | 99.7 | 621 | 14 | US-10-022-249-49 | Sequence 49, Appl |
| 17 | 3303 | 99.7 | 621 | 14 | US-10-022-249-51 | Sequence 51, Appl |
| 18 | 3303 | 99.7 | 621 | 14 | US-10-022-249-153 | Sequence 153, App |
| 19 | 3303 | 99.7 | 621 | 14 | US-10-022-249-157 | Sequence 157, App |
| 20 | 3303 | 99.7 | 621 | 14 | US-10-022-249-169 | Sequence 169, App |
| 21 | 3303 | 99.7 | 621 | 14 | US-10-022-249-245 | Sequence 245, App |
| 22 | 3303 | 99.7 | 621 | 14 | US-10-022-249-253 | Sequence 253, App |
| 23 | 3303 | 99.7 | 621 | 14 | US-10-022-249-257 | Sequence 257, App |
| 24 | 3303 | 99.7 | 621 | 14 | US-10-022-249-261 | Sequence 261, App |
| 25 | 3303 | 99.7 | 621 | 14 | US-10-022-249-269 | Sequence 269, App |
| 26 | 3303 | 99.7 | 621 | 14 | US-10-022-249-281 | Sequence 281, App |
| 27 | 3302 | 99.7 | 621 | 14 | US-10-022-390-47 | Sequence 47, Appl |
| 28 | 3302 | 99.7 | 621 | 14 | US-10-022-390-53 | Sequence 53, Appl |
| 29 | 3302 | 99.7 | 621 | 14 | US-10-022-390-137 | Sequence 137, App |
| 30 | 3302 | 99.7 | 621 | 14 | US-10-022-390-265 | Sequence 265, App |
| 31 | 3302 | 99.7 | 621 | 14 | US-10-022-390-295 | Sequence 295, App |
| 32 | 3302 | 99.7 | 621 | 14 | US-10-022-249-47 | Sequence 47, Appl |
| 33 | 3302 | 99.7 | 621 | 14 | US-10-022-249-53 | Sequence 53, Appl |
| 34 | 3302 | 99.7 | 621 | 14 | US-10-022-249-137 | Sequence 137, App |
| 35 | 3302 | 99.7 | 621 | 14 | US-10-022-249-265 | Sequence 265, App |
| 36 | 3302 | 99.7 | 621 | 14 | US-10-022-249-295 | Sequence 295, App |
| 37 | 3301 | 99.7 | 621 | 14 | US-10-022-390-7 | Sequence 7, Appl1 |
| 38 | 3301 | 99.7 | 621 | 14 | US-10-022-390-15 | Sequence 15, Appl |
| 39 | 3301 | 99.7 | 621 | 14 | US-10-022-390-17 | Sequence 17, Appl |
| 40 | 3301 | 99.7 | 621 | 14 | US-10-022-390-19 | Sequence 19, Appl |
| 41 | 3301 | 99.7 | 621 | 14 | US-10-022-390-29 | Sequence 29, Appl |
| 42 | 3301 | 99.7 | 621 | 14 | US-10-022-390-33 | Sequence 33, Appl |
| 43 | 3301 | 99.7 | 621 | 14 | US-10-022-390-39 | Sequence 39, Appl |
| 44 | 3301 | 99.7 | 621 | 14 | US-10-022-390-45 | Sequence 45, Appl |
| 45 | 3301 | 99.7 | 621 | 14 | US-10-022-390-57 | Sequence 57, Appl |

ALIGNMENTS

RESULT 1
US-10-022-390-113
; Sequence 113, Application US/10022390
; Publication No. US20030129203A1
; GENERAL INFORMATION:
; APPLICANT: Vega, Manuel
; APPLICANT: Dittanti, Lila
; APPLICANT: Flaux, Marjorie
; TITLE OF INVENTION: MUTANT RECOMBINANT ADENO-ASSOCIATED VIRUSES
; FILE REFERENCE: 37851-912
; CURRENT APPLICATION NUMBER: US/10/022,390
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/315,382
; PRIOR FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant rep protein: rep 78 350 AAT
US-10-022-390-113

| | | | | |
|-----------------------|-----------------|---|-----------|---|
| Query Match | 100.0%; | Score 3312; | DB 14; | Length 621; |
| Best Local Similarity | 100.0%; | Pred. No. 3.8e-285; | | |
| Matches 621; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY | 1 | TAGFYEIVIKVPSDLDEHLP | 1 | TAGFYEIVIKVPSDLDEHLP |
| Db | 1 | TAGFYEIVIKVPSDLDEHLP | 1 | TAGFYEIVIKVPSDLDEHLP |
| QY | 61 | RDFTLEWRRVSKAPEALFFVQFEKGESYFHMVLTGKSMVLGRFLSQIREKLIQRI | 120 | RDFTLEWRRVSKAPEALFFVQFEKGESYFHMVLTGKSMVLGRFLSQIREKLIQRI |
| Db | 61 | RDFTLEWRRVSKAPEALFFVQFEKGESYFHMVLTGKSMVLGRFLSQIREKLIQRI | 120 | RDFTLEWRRVSKAPEALFFVQFEKGESYFHMVLTGKSMVLGRFLSQIREKLIQRI |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2005, 04:50:54 ; Search time 42 Seconds
(without alignments)
1103.740 Million cell updates/sec

Title: US-10-022-390-113
Perfect score: 3312
Sequence: 1 TAGFYEIVIKVPSDLDEHLP.....DACTACDLNVNVDLDCIFEQ 621

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 3034 | 91.6 | 623 | 4 | US-09-532-594B-2 |
| 2 | 3034 | 91.6 | 623 | 4 | US-09-532-594B-11 |
| 3 | 2954 | 89.2 | 623 | 4 | US-09-807-802A-2 |
| 4 | 2954 | 89.2 | 623 | 4 | US-09-807-802A-5 |
| 5 | 2618 | 79.0 | 536 | 4 | US-09-532-594B-10 |
| 6 | 2582.5 | 78.0 | 546 | 4 | US-09-807-802A-7 |
| 7 | 1926 | 58.2 | 399 | 4 | US-09-532-594B-9 |
| 8 | 1823 | 55.0 | 399 | 4 | US-09-807-802A-9 |
| 9 | 1510 | 45.6 | 312 | 4 | US-09-532-594B-8 |
| 10 | 1451.5 | 43.8 | 322 | 4 | US-09-807-802A-11 |
| 11 | 347.5 | 10.5 | 276 | 3 | US-09-171-461-3 |
| 12 | 347.5 | 10.5 | 276 | 4 | US-09-970-711-3 |
| 13 | 135 | 4.1 | 644 | 3 | US-09-300-909-18 |
| 14 | 126.5 | 3.8 | 592 | 4 | US-09-861-451A-72 |
| 15 | 125 | 3.8 | 605 | 2 | US-08-472-666-1 |
| 16 | 125 | 3.8 | 605 | 5 | PCT-US96-07615-1 |
| 17 | 115 | 3.5 | 543 | 2 | US-08-823-516-143 |
| 18 | 115 | 3.5 | 543 | 4 | US-09-940-244-143 |
| 19 | 115 | 3.5 | 647 | 3 | US-09-300-909-17 |
| 20 | 114 | 3.4 | 872 | 4 | US-09-198-452A-163 |
| 21 | 114 | 3.4 | 873 | 4 | US-09-438-185A-146 |
| 22 | 110 | 3.3 | 937 | 1 | US-08-253-155A-31 |
| 23 | 110 | 3.3 | 937 | 4 | US-09-538-092-1092 |
| 24 | 110 | 3.3 | 959 | 4 | US-09-538-092-1091 |
| 25 | 109.5 | 3.3 | 662 | 4 | US-09-107-532A-5989 |
| 26 | 109 | 3.3 | 646 | 3 | US-09-300-909-14 |
| 27 | 107.5 | 3.2 | 657 | 3 | US-09-300-909-16 |

| | | | | | | |
|----|-------|-----|-------|---|----------------------|--------------------|
| 28 | 107 | 3.2 | 719 | 4 | US-09-949-016-7766 | Sequence 7766, Ap |
| 29 | 105.5 | 3.2 | 1092 | 4 | US-09-543-681A-7058 | Sequence 7058, Ap |
| 30 | 105.5 | 3.2 | 1095 | 4 | US-09-710-279-3154 | Sequence 3154, Ap |
| 31 | 105.5 | 3.2 | 10182 | 3 | US-09-134-001C-3159 | Sequence 3159, Ap |
| 32 | 105 | 3.2 | 672 | 3 | US-09-040-843-4 | Sequence 4, Appli |
| 33 | 105 | 3.2 | 672 | 3 | US-09-621-855-4 | Sequence 4, Appli |
| 34 | 105 | 3.2 | 809 | 4 | US-09-252-991A-29280 | Sequence 29280, A |
| 35 | 105 | 3.2 | 866 | 3 | US-09-040-843-2 | Sequence 2, Appli |
| 36 | 105 | 3.2 | 866 | 3 | US-09-621-855-2 | Sequence 2, Appli |
| 37 | 104 | 3.1 | 823 | 4 | US-09-107-532A-5667 | Sequence 5667, Ap |
| 38 | 102.5 | 3.1 | 339 | 3 | US-09-171-461-36 | Sequence 36, Appli |
| 39 | 102.5 | 3.1 | 339 | 4 | US-09-970-711-36 | Sequence 36, Appli |
| 40 | 102.5 | 3.1 | 506 | 4 | US-09-902-540-14235 | Sequence 14235, A |
| 41 | 102 | 3.1 | 565 | 4 | US-09-543-681A-4569 | Sequence 4569, Ap |
| 42 | 102 | 3.1 | 724 | 4 | US-09-489-039A-12100 | Sequence 12100, A |
| 43 | 102 | 3.1 | 952 | 4 | US-09-248-796A-20878 | Sequence 20878, A |
| 44 | 102 | 3.1 | 3077 | 6 | 5223423-2 | Patent No. 5223423 |
| 45 | 102 | 3.1 | 3077 | 6 | 5223423-2 | Patent No. 5223423 |

ALIGNMENTS

RESULT 1

US-09-532-594B-2
; Sequence 2, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 Rep protein (full length)
US-09-532-594B-2

| | | | | |
|-----------------------|--------|---|--|---|
| Query Match | -91.6% | Score 3034; | DB 4; | Length 623; |
| Best Local Similarity | 90.3% | Pred. No. 2.1e-303; | | |
| Matches | 561; | Conservative | 31; | Mismatches 27; Indels 2; Gaps 2; |
| Qy | 3 | GPYEIVIKVPSDLDEHLP | IGISDSFVNWVAEKEWELPPDSMDLNLIEQAPLTVAEKLQRD | 62 |
| Db | 3 | GPYEIVLKVP | SDLEHLP | IGISDSFVSWVAEKEWELPPDSMDLNLIEQAPLTVAEKLQRE 62 |
| Qy | 63 | FLTEWRRVSKAPEALFFVQPEKGSYFHMVHLVETTVGKSMVLGRFLSQIREKLIQRIYR | 122 | |
| Db | 63 | FLVEWRRVSKAPEALFFVQPEKGSYFHLHLVETTVGKSMVVGRYVSVQIKEKLVTRIYR | 122 | |
| Qy | 123 | GIEPTLPNWFVAVTKTRNGAGGKNKVVDECYIPNYLLPKTQPELQWANTNMEQYLSACLNL | 182 | |
| Db | 123 | GVEPQLPNWFVAVTKTRNGAGGKNKVVDDCYIPNYLLPKTQPELQWANTNMDQYISACLNL | 182 | |
| Qy | 183 | TERKRLVAQHLTHVSQTQEQNKENQNPNSDAPVIRKTSARYMELVGLVDKGITSEKQW | 242 | |
| Db | 183 | ABERKRLVAQHLTHVSQTQEQNKENQNPNSDAPVIRKTSARYMELVGLVDRGITSEKQW | 242 | |
| Qy | 243 | IQEDQASYISFNAASRSQIKAALDNAGKIMSLTKTAPDYLVGQOPVEDISSNRIYKIL | 302 | |
| Db | 243 | IQEDQASYISFNAASRSQIKAALDNASKIMSLTKTAPDYLVGQNPVEDISSNRIYRIL | 302 | |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: February 23, 2005, 01:53:32 ; Search time 7486 Seconds
(without alignments)
4019.595 Million cell updates/sec

Title: US-10-022-390-113
Perfect score: 3312
Sequence: 1 TAGFYRIVIKVPSDLDEHLP.....DACTACDLVNVLDLDCIFEQ 621

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_l/USPTO_spool/US10022390/runat_17022005_095009_14654/app_query.fasta_1.775
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10022390 @CGN_1_1_4694 @runat_17022005_095009_14654 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb_ba:
2: gb_htg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_ats:
12: gb_sy:
13: gb_un:
14: gb_vi:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|-------------------|
| 1 | 3312 | 100.0 | 1932 | 6 AX712816 | AX712816 Sequence |
| 2 | 3312 | 100.0 | 1932 | 6 AX719925 | AX719925 Sequence |
| 3 | 3306 | 99.8 | 1932 | 6 AX712941 | AX712941 Sequence |
| 4 | 3306 | 99.8 | 1932 | 6 AX712942 | AX712942 Sequence |

| | | | | | | |
|----|------|------|------|---|----------|-------------------|
| 5 | 3306 | 99.8 | 1932 | 6 | AX712943 | AX712943 Sequence |
| 6 | 3306 | 99.8 | 1932 | 6 | AX712944 | AX712944 Sequence |
| 7 | 3306 | 99.8 | 1932 | 6 | AX720050 | AX720050 Sequence |
| 8 | 3306 | 99.8 | 1932 | 6 | AX720051 | AX720051 Sequence |
| 9 | 3306 | 99.8 | 1932 | 6 | AX720052 | AX720052 Sequence |
| 10 | 3306 | 99.8 | 1932 | 6 | AX720053 | AX720053 Sequence |
| 11 | 3304 | 99.8 | 1932 | 6 | AX712817 | AX712817 Sequence |
| 12 | 3304 | 99.8 | 1932 | 6 | AX719926 | AX719926 Sequence |
| 13 | 3303 | 99.7 | 1932 | 6 | AX712799 | AX712799 Sequence |
| 14 | 3303 | 99.7 | 1932 | 6 | AX712800 | AX712800 Sequence |
| 15 | 3303 | 99.7 | 1932 | 6 | AX712826 | AX712826 Sequence |
| 16 | 3303 | 99.7 | 1932 | 6 | AX712827 | AX712827 Sequence |
| 17 | 3303 | 99.7 | 1932 | 6 | AX712830 | AX712830 Sequence |
| 18 | 3303 | 99.7 | 1932 | 6 | AX712849 | AX712849 Sequence |
| 19 | 3303 | 99.7 | 1932 | 6 | AX712851 | AX712851 Sequence |
| 20 | 3303 | 99.7 | 1932 | 6 | AX712853 | AX712853 Sequence |
| 21 | 3303 | 99.7 | 1932 | 6 | AX712855 | AX712855 Sequence |
| 22 | 3303 | 99.7 | 1932 | 6 | AX712858 | AX712858 Sequence |
| 23 | 3303 | 99.7 | 1932 | 6 | AX719908 | AX719908 Sequence |
| 24 | 3303 | 99.7 | 1932 | 6 | AX719909 | AX719909 Sequence |
| 25 | 3303 | 99.7 | 1932 | 6 | AX719935 | AX719935 Sequence |
| 26 | 3303 | 99.7 | 1932 | 6 | AX719936 | AX719936 Sequence |
| 27 | 3303 | 99.7 | 1932 | 6 | AX719939 | AX719939 Sequence |
| 28 | 3303 | 99.7 | 1932 | 6 | AX719958 | AX719958 Sequence |
| 29 | 3303 | 99.7 | 1932 | 6 | AX719960 | AX719960 Sequence |
| 30 | 3303 | 99.7 | 1932 | 6 | AX719962 | AX719962 Sequence |
| 31 | 3303 | 99.7 | 1932 | 6 | AX719964 | AX719964 Sequence |
| 32 | 3303 | 99.7 | 1932 | 6 | AX719967 | AX719967 Sequence |
| 33 | 3302 | 99.7 | 1932 | 6 | AX712798 | AX712798 Sequence |
| 34 | 3302 | 99.7 | 1932 | 6 | AX712801 | AX712801 Sequence |
| 35 | 3302 | 99.7 | 1932 | 6 | AX712822 | AX712822 Sequence |
| 36 | 3302 | 99.7 | 1932 | 6 | AX712854 | AX712854 Sequence |
| 37 | 3302 | 99.7 | 1932 | 6 | AX712861 | AX712861 Sequence |
| 38 | 3302 | 99.7 | 1932 | 6 | AX719907 | AX719907 Sequence |
| 39 | 3302 | 99.7 | 1932 | 6 | AX719910 | AX719910 Sequence |
| 40 | 3302 | 99.7 | 1932 | 6 | AX719931 | AX719931 Sequence |
| 41 | 3302 | 99.7 | 1932 | 6 | AX719963 | AX719963 Sequence |
| 42 | 3302 | 99.7 | 1932 | 6 | AX719970 | AX719970 Sequence |
| 43 | 3301 | 99.7 | 1932 | 6 | AX712778 | AX712778 Sequence |
| 44 | 3301 | 99.7 | 1932 | 6 | AX712782 | AX712782 Sequence |
| 45 | 3301 | 99.7 | 1932 | 6 | AX712783 | AX712783 Sequence |

ALIGNMENTS

RESULT 1
AX712816
LOCUS AX712816 1932 bp DNA linear PAT 11-APR-2003
DEFINITION Sequence 604 from Patent WO03018820.
ACCESSION AX712816
VERSION AX712816.1 GI:29823475
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Vega,M., Drittanti,L. and Plaux,M.
TITLE Mutant recombinant adeno-associated viruses related applications
JOURNAL Patent: WO 03018820-A 604 06-MAR-2003;
Nautilus Biotech (FR)
FEATURES Location/Qualifiers
source 1..1932
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Mutant rep DNA sequence: 350 AAT"

ORIGIN
Alignment Scores:
Pred. No.: 2.44e-270 Length: 1932
Score: 3312.00 Matches: 621
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 23, 2005, 01:42:38 ; Search time 894 Seconds
(without alignments)
4112.035 Million cell updates/sec

Title: US-10-022-390-113
Perfect score: 3312
Sequence: 1 TAGFYEIVIKVPSDLDEHLP.....DACTACDLNVNVDLDDCIFEQ 621

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10022390/runat_17022005_095008_14644/app_query.fasta_1.775
-DB=N Geneseq_16Dec04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10022390 @CGN_1_1_644 @runat_17022005_095008_14644 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 3312 | 100.0 | 1932 | 8 ACD01626 | AcD01626 DNA seque |
| 2 | 3312 | 100.0 | 1932 | 10 ADB71023 | Adb71023 Adeno-ass |
| 3 | 3306 | 99.8 | 1932 | 8 ACD01752 | AcD01752 High thro |
| 4 | 3306 | 99.8 | 1932 | 8 ACD01753 | AcD01753 High thro |
| 5 | 3306 | 99.8 | 1932 | 8 ACD01751 | AcD01751 High thro |

| | | | | | |
|----|------|------|------|-------------|--------------------|
| 6 | 3306 | 99.8 | 1932 | 8 ACD01754 | AcD01754 High thro |
| 7 | 3306 | 99.8 | 1932 | 10 ADB71148 | Adb71148 Adeno-ass |
| 8 | 3306 | 99.8 | 1932 | 10 ADB71151 | Adb71151 Adeno-ass |
| 9 | 3306 | 99.8 | 1932 | 10 ADB71149 | Adb71149 Adeno-ass |
| 10 | 3306 | 99.8 | 1932 | 10 ADB71150 | Adb71150 Adeno-ass |
| 11 | 3304 | 99.8 | 1932 | 8 ACD01627 | AcD01627 DNA seque |
| 12 | 3304 | 99.8 | 1932 | 10 ADB71024 | Adb71024 Adeno-ass |
| 13 | 3303 | 99.7 | 1932 | 8 ACD01637 | AcD01637 DNA seque |
| 14 | 3303 | 99.7 | 1932 | 8 ACD01636 | AcD01636 DNA seque |
| 15 | 3303 | 99.7 | 1932 | 8 ACD01610 | AcD01610 DNA seque |
| 16 | 3303 | 99.7 | 1932 | 8 ACD01640 | AcD01640 DNA seque |
| 17 | 3303 | 99.7 | 1932 | 8 ACD01609 | AcD01609 DNA seque |
| 18 | 3303 | 99.7 | 1932 | 8 ACD01659 | AcD01659 DNA seque |
| 19 | 3303 | 99.7 | 1932 | 8 ACD01665 | AcD01665 DNA seque |
| 20 | 3303 | 99.7 | 1932 | 8 ACD01661 | AcD01661 DNA seque |
| 21 | 3303 | 99.7 | 1932 | 8 ACD01668 | AcD01668 DNA seque |
| 22 | 3303 | 99.7 | 1932 | 8 ACD01663 | AcD01663 DNA seque |
| 23 | 3303 | 99.7 | 1932 | 10 ADB71006 | Adb71006 Adeno-ass |
| 24 | 3303 | 99.7 | 1932 | 10 ADB71056 | Adb71056 Adeno-ass |
| 25 | 3303 | 99.7 | 1932 | 10 ADB71065 | Adb71065 Adeno-ass |
| 26 | 3303 | 99.7 | 1932 | 10 ADB71033 | Adb71033 Adeno-ass |
| 27 | 3303 | 99.7 | 1932 | 10 ADB71058 | Adb71058 Adeno-ass |
| 28 | 3303 | 99.7 | 1932 | 10 ADB71007 | Adb71007 Adeno-ass |
| 29 | 3303 | 99.7 | 1932 | 10 ADB71062 | Adb71062 Adeno-ass |
| 30 | 3303 | 99.7 | 1932 | 10 ADB71037 | Adb71037 Adeno-ass |
| 31 | 3303 | 99.7 | 1932 | 10 ADB71060 | Adb71060 Adeno-ass |
| 32 | 3303 | 99.7 | 1932 | 10 ADB71034 | Adb71034 Adeno-ass |
| 33 | 3302 | 99.7 | 1932 | 8 ACD01671 | AcD01671 DNA seque |
| 34 | 3302 | 99.7 | 1932 | 8 ACD01611 | AcD01611 DNA seque |
| 35 | 3302 | 99.7 | 1932 | 8 ACD01632 | AcD01632 DNA seque |
| 36 | 3302 | 99.7 | 1932 | 8 ACD01664 | AcD01664 DNA seque |
| 37 | 3302 | 99.7 | 1932 | 8 ACD01608 | AcD01608 DNA seque |
| 38 | 3302 | 99.7 | 1932 | 10 ADB71061 | Adb71061 Adeno-ass |
| 39 | 3302 | 99.7 | 1932 | 10 ADB71005 | Adb71005 Adeno-ass |
| 40 | 3302 | 99.7 | 1932 | 10 ADB71029 | Adb71029 Adeno-ass |
| 41 | 3302 | 99.7 | 1932 | 10 ADB71008 | Adb71008 Adeno-ass |
| 42 | 3302 | 99.7 | 1932 | 10 ADB71068 | Adb71068 Adeno-ass |
| 43 | 3301 | 99.7 | 1932 | 8 ACD01625 | AcD01625 DNA seque |
| 44 | 3301 | 99.7 | 1932 | 8 ACD01607 | AcD01607 DNA seque |
| 45 | 3301 | 99.7 | 1932 | 8 ACD01594 | AcD01594 DNA seque |

ALIGNMENTS

RESULT 1
ACD01626
ID ACD01626 standard; DNA; 1932 BP.
XX
AC ACD01626;
XX
DT 30-JUL-2003 (first entry)
XX
DE DNA sequence #42 encoding Adeno associated virus mutant rep protein.
XX
KW Polypeptide production; protein production; target protein;
KW high throughput directed evolution; rational mutagenesis; AAV;
KW protein variant generation; virus titring; Adeno associated virus;
KW rep protein; mutant; ds.
XX
OS Adeno associated virus.
OS Synthetic.
XX
PN WO2003023032-A2.
XX
PD 20-MAR-2003.
XX
PF 16-AUG-2002; 2002WO-IB003921.
XX
PR 27-AUG-2001; 2001US-0315382P.
PR 17-DEC-2001; 2001US-00022249.
XX
PA (NAUT-) NAUTILUS BIOTECH.
XX

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: February 23, 2005, 06:05:03 ; Search time 292 Seconds
(without alignments)
3479.890 Million cell updates/sec

Title: US-10-022-390-113
 Perfect score: 3312
 Sequence: 1 TAGFYEIVIKVPSLDLDEHLP.....DACTACDLVNVDLDDCIFEQ 621

| Scoring table: | |
|----------------|-------------|
| BLOSUM62 | |
| Xgapop 10.0 , | Xgapext 0.5 |
| Ygapop 10.0 , | Ygapext 0.5 |
| Fgapop 6.0 , | Fgapext 7.0 |
| Delop .6.0 , | Delext .7.0 |

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 24055568

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10022390/runat_17022005_095011_14680/app_query.fasta_1.775
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10022390@cGN_1_1_177 @runat_17022005_095011_14680 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

```
Database : Issued Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-------------------|--------------------|
| 1 | 3297 | 99.5 | 7557 | 3 | US-09-770-315-3 | Sequence 3, Appli |
| 2 | 3297 | 99.5 | 8698 | 3 | US-09-770-315-2 | Sequence 2, Appli |
| 3 | 3290 | 99.3 | 4680 | 1 | US-08-254-358-1 | Sequence 1, Appli |
| 4 | 3290 | 99.3 | 4680 | 1 | US-08-475-391-1 | Sequence 1, Appli |
| 5 | 3290 | 99.3 | 4680 | 2 | US-08-709-609-1 | Sequence 1, Appli |
| 6 | 3290 | 99.3 | 4680 | 5 | PCT-US95-07178-1 | Sequence 1, Appli |
| 7 | 3290 | 99.3 | 4681 | 4 | US-09-807-802A-18 | Sequence 18, Appli |
| C 8 | 3290 | 99.3 | 4910 | 2 | US-08-331-384-2 | Sequence 2, Appli |
| C 9 | 3290 | 99.3 | 4910 | 2 | US-08-836-087-2 | Sequence 2, Appli |
| C 10 | 3290 | 99.3 | 4910 | 3 | US-09-246-320-2 | Sequence 2, Appli |
| C 11 | 3290 | 99.3 | 4910 | 3 | US-09-546-738-2 | Sequence 2, Appli |
| 12 | 3290 | 99.3 | 8151 | 4 | US-09-438-268-2 | Sequence 2, Appli |

| | | | | | | | |
|---|-----|--------|------|---------|---|-------------------|--------------------|
| C | 13 | 3290 | 99.3 | 8179 | 4 | US-09-438-268-5 | Sequence 5, Appli |
| | 14 | 3181 | 96.0 | 7214 | 4 | US-09-438-268-1 | Sequence 1, Appli |
| | 15 | 3034 | 91.6 | 1872 | 3 | US-09-532-594B-3 | Sequence 3, Appli |
| | 16 | 3034 | 91.6 | 1872 | 3 | US-09-532-594B-15 | Sequence 15, Appli |
| | 17 | 3034 | 91.6 | 4767 | 3 | US-09-532-594B-1 | Sequence 1, Appli |
| | 18 | 2954 | 89.2 | 1872 | 4 | US-09-807-802A-4 | Sequence 4, Appli |
| | 19 | 2954 | 89.2 | 4718 | 4 | US-09-807-802A-1 | Sequence 1, Appli |
| | 20 | 2936 | 88.6 | 4683 | 4 | US-09-807-802A-19 | Sequence 19, Appli |
| | 21 | 2618 | 79.0 | 1611 | 3 | US-09-532-594B-14 | Sequence 14, Appli |
| | 22 | 2582.5 | 78.0 | 1641 | 4 | US-09-807-802A-6 | Sequence 6, Appli |
| | 23 | 1926 | 58.2 | 1197 | 3 | US-09-532-594B-13 | Sequence 13, Appli |
| | 24 | 1823 | 55.0 | 1200 | 4 | US-09-807-802A-8 | Sequence 8, Appli |
| | 25 | 1510 | 45.6 | 939 | 3 | US-09-532-594B-12 | Sequence 12, Appli |
| | 26. | 1451.5 | 43.8 | 969 | 4 | US-09-807-802A-10 | Sequence 10, Appli |
| | 27 | 566 | 17.1 | 4072 | 3 | US-09-770-315-4 | Sequence 4, Appli |
| | 28 | 443.5 | 13.4 | 5049 | 1 | US-08-336-345-1 | Sequence 1, Appli |
| | 29 | 443.5 | 13.4 | 5049 | 1 | US-08-336-345-2 | Sequence 2, Appli |
| | 30 | 443.5 | 13.4 | 5049 | 2 | US-08-647-655-1 | Sequence 1, Appli |
| | 31 | 443.5 | 13.4 | 5049 | 2 | US-08-647-655-2 | Sequence 2, Appli |
| | 32 | 351.5 | 10.6 | 43804 | 3 | US-09-171-461-1 | Sequence 1, Appli |
| C | 33 | 351.5 | 10.6 | 43804 | 4 | US-09-970-711-1 | Sequence 1, Appli |
| | 34 | 290 | 8.8 | 349 | 1 | US-08-510-790-4 | Sequence 4, Appli |
| | 35 | 141.5 | 4.3 | 5910 | 1 | US-08-195-814-1 | Sequence 1, Appli |
| | 36 | 126.5 | 3.8 | 8106 | 3 | US-09-135-241-1 | Sequence 1, Appli |
| | 37 | 125.5 | 3.8 | 1776 | 4 | US-09-861-451A-71 | Sequence 71, Appli |
| | 38 | 123 | 3.7 | 43804 | 3 | US-09-171-461-1 | Sequence 1, Appli |
| | 39 | 123 | 3.7 | 43804 | 4 | US-09-970-711-1 | Sequence 1, Appli |
| | 40 | 115.5 | 3.5 | 5152 | 3 | US-09-690-364-10 | Sequence 10, Appli |
| | 41 | 115 | 3.5 | 7833 | 1 | US-08-074-879-9 | Sequence 9, Appli |
| | 42 | 115 | 3.5 | 7833 | 1 | US-08-468-057A-9 | Sequence 9, Appli |
| C | 43 | 114 | 3.4 | 31096 | 4 | US-08-956-171E-59 | Sequence 59, Appli |
| | 44 | 114 | 3.4 | 31096 | 4 | US-08-781-986A-59 | Sequence 59, Appli |
| | 45 | 114 | 3.4 | 1230025 | 4 | US-09-198-452A-1 | Sequence 1, Appli |

ALIGNMENTS

```

RESULT 1
US-09-770-315-3
; Sequence 3, Application US/09770315
; Patent No. 6429001
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,536
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 7557
;   TYPE: DNA
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: recombinant DNA
US-09-770-315-3

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 23, 2005, 09:34:18 ; Search time 925 Seconds
(without alignments)
3968.020 Million cell updates/sec

Title: US-10-022-390-113
Perfect score: 3312
Sequence: 1 TAGFYRIVIKVPSDLDEHLP.....DACTACDLNVNVDLDCIFEQ 621

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5384158 seqs, 2955248155 residues
Total number of hits satisfying chosen parameters: 10768316

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10022390/runat 17022005 095013 14760/app query.fasta 1.775
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10022390 @CGN 1 1 678 @runat 17022005 095013 14760
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-------------------|-------------------|
| 1 | 3312 | 100.0 | 1932 | 15 | US-10-022-390-604 | Sequence 604, App |
| 2 | 3312 | 100.0 | 1932 | 15 | US-10-022-249-604 | Sequence 604, App |
| 3 | 3306 | 99.8 | 1932 | 15 | US-10-022-390-729 | Sequence 729, App |
| 4 | 3306 | 99.8 | 1932 | 15 | US-10-022-390-730 | Sequence 730, App |
| 5 | 3306 | 99.8 | 1932 | 15 | US-10-022-390-731 | Sequence 731, App |
| 6 | 3306 | 99.8 | 1932 | 15 | US-10-022-390-732 | Sequence 732, App |
| 7 | 3306 | 99.8 | 1932 | 15 | US-10-022-249-729 | Sequence 729, App |
| 8 | 3306 | 99.8 | 1932 | 15 | US-10-022-249-730 | Sequence 730, App |
| 9 | 3306 | 99.8 | 1932 | 15 | US-10-022-249-731 | Sequence 731, App |
| 10 | 3306 | 99.8 | 1932 | 15 | US-10-022-249-732 | Sequence 732, App |
| 11 | 3304 | 99.8 | 1932 | 15 | US-10-022-390-605 | Sequence 605, App |
| 12 | 3304 | 99.8 | 1932 | 15 | US-10-022-249-605 | Sequence 605, App |
| 13 | 3303 | 99.7 | 1932 | 15 | US-10-022-390-587 | Sequence 587, App |
| 14 | 3303 | 99.7 | 1932 | 15 | US-10-022-390-588 | Sequence 588, App |
| 15 | 3303 | 99.7 | 1932 | 15 | US-10-022-390-614 | Sequence 614, App |
| 16 | 3303 | 99.7 | 1932 | 15 | US-10-022-390-615 | Sequence 615, App |
| 17 | 3303 | 99.7 | 1932 | 15 | US-10-022-390-618 | Sequence 618, App |
| 18 | 3303 | 99.7 | 1932 | 15 | US-10-022-390-637 | Sequence 637, App |
| 19 | 3303 | 99.7 | 1932 | 15 | US-10-022-390-639 | Sequence 639, App |
| 20 | 3303 | 99.7 | 1932 | 15 | US-10-022-390-641 | Sequence 641, App |
| 21 | 3303 | 99.7 | 1932 | 15 | US-10-022-390-643 | Sequence 643, App |
| 22 | 3303 | 99.7 | 1932 | 15 | US-10-022-390-646 | Sequence 646, App |
| 23 | 3303 | 99.7 | 1932 | 15 | US-10-022-249-587 | Sequence 587, App |
| 24 | 3303 | 99.7 | 1932 | 15 | US-10-022-249-588 | Sequence 588, App |
| 25 | 3303 | 99.7 | 1932 | 15 | US-10-022-249-614 | Sequence 614, App |
| 26 | 3303 | 99.7 | 1932 | 15 | US-10-022-249-615 | Sequence 615, App |
| 27 | 3303 | 99.7 | 1932 | 15 | US-10-022-249-618 | Sequence 618, App |
| 28 | 3303 | 99.7 | 1932 | 15 | US-10-022-249-637 | Sequence 637, App |
| 29 | 3303 | 99.7 | 1932 | 15 | US-10-022-249-639 | Sequence 639, App |
| 30 | 3303 | 99.7 | 1932 | 15 | US-10-022-249-641 | Sequence 641, App |
| 31 | 3303 | 99.7 | 1932 | 15 | US-10-022-249-643 | Sequence 643, App |
| 32 | 3303 | 99.7 | 1932 | 15 | US-10-022-249-646 | Sequence 646, App |
| 33 | 3302 | 99.7 | 1932 | 15 | US-10-022-390-586 | Sequence 586, App |
| 34 | 3302 | 99.7 | 1932 | 15 | US-10-022-390-589 | Sequence 589, App |
| 35 | 3302 | 99.7 | 1932 | 15 | US-10-022-390-610 | Sequence 610, App |
| 36 | 3302 | 99.7 | 1932 | 15 | US-10-022-390-642 | Sequence 642, App |
| 37 | 3302 | 99.7 | 1932 | 15 | US-10-022-390-649 | Sequence 649, App |
| 38 | 3302 | 99.7 | 1932 | 15 | US-10-022-249-586 | Sequence 586, App |
| 39 | 3302 | 99.7 | 1932 | 15 | US-10-022-249-589 | Sequence 589, App |
| 40 | 3302 | 99.7 | 1932 | 15 | US-10-022-249-610 | Sequence 610, App |
| 41 | 3302 | 99.7 | 1932 | 15 | US-10-022-249-642 | Sequence 642, App |
| 42 | 3302 | 99.7 | 1932 | 15 | US-10-022-249-649 | Sequence 649, App |
| 43 | 3301 | 99.7 | 1932 | 15 | US-10-022-390-566 | Sequence 566, App |
| 44 | 3301 | 99.7 | 1932 | 15 | US-10-022-390-570 | Sequence 570, App |
| 45 | 3301 | 99.7 | 1932 | 15 | US-10-022-390-571 | Sequence 571, App |

ALIGNMENTS

RESULT 1
US-10-022-390-604
; Sequence 604, Application US/10022390
; Publication No. US20030129203A1
; GENERAL INFORMATION:
; APPLICANT: Vega, Manuel
; APPLICANT: Dittanti, Lila
; APPLICANT: Flaux, Marjorie
; TITLE OF INVENTION: MUTANT RECOMBINANT ADENO-ASSOCIATED VIRUSES
; FILE REFERENCE: 37851-912
; CURRENT APPLICATION NUMBER: US/10/022,390
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/315,382
; PRIOR FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 604
; LENGTH: 1932
; TYPE: DNA
; ORGANISM: Artificial Sequence

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.
OM protein - nucleic search, using frame_plus_p2n model
Run on: February 23, 2005, 06:22:53 ; Search time 5412 Seconds
(without alignments)
4367.685 Million cell updates/sec

Title: US-10-022-390-113
Perfect score: 3312
Sequence: 1 TAGFYEIVIKVPSDLDEHLP.....DACTACDLNVNVDLDCIFEQ 621

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10022390/runat_17022005_095010_14666/app_query.fasta_1.775
-DB=EST -QFMT=fastap -SUFFIX=rat -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10022390@cgn_1_1_4352@runat_17022005_095010_14666 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hsc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | | | |
|------------|-------|--------------------------|--------------------|
| Result No. | Score | Query Match Length DB ID | Description |
| 1 | 264.5 | 8.0 489 8 | AZ289042 RPCI-23-5 |
| C 2 | 252.5 | 7.6 778 9 | AG491391 Mus muscu |
| C 3 | 249.5 | 7.5 495 8 | AZ028362 RPCI-23-3 |
| C 4 | 159.5 | 4.8 484 9 | CE239279 tigr-gss- |
| 5 | 124 | 3.7 585 6 | CD793420 EST664781 |
| 6 | 122.5 | 3.7 3053 3 | CNS0AAM3 Arabidops |
| 7 | 119.5 | 3.6 3267 9 | CL961562 OsIFCC006 |
| 8 | 114 | 3.4 3758 3 | AK030099 Mus muscu |
| 9 | 113 | 3.4 3546 3 | BC033413 Mus muscu |

| | | | | | |
|------|-------|-----|------|---|--------------------|
| 10 | 112.5 | 3.4 | 1140 | 1 | AF122105 |
| 11 | 112.5 | 3.4 | 2398 | 3 | BC027369 Mus muscu |
| 12 | 112.5 | 3.4 | 2556 | 9 | CL974261 OsIFCC025 |
| 13 | 112 | 3.4 | 6448 | 3 | BC083191 Mus muscu |
| 14 | 111 | 3.4 | 3042 | 3 | CR603800 full-leng |
| 15 | 110.5 | 3.3 | 2501 | 3 | AK033631 Mus muscu |
| 16 | 110.5 | 3.3 | 2808 | 9 | CL966666 OsIFCC013 |
| 17 | 110.5 | 3.3 | 2912 | 3 | AK030041 Mus muscu |
| 18 | 110.5 | 3.3 | 2966 | 3 | BC050106 Mus muscu |
| 19 | 110.5 | 3.3 | 3380 | 3 | BC040795 Mus muscu |
| 20 | 110.5 | 3.3 | 3823 | 3 | AK035273 Mus muscu |
| 21 | 110 | 3.3 | 751 | 1 | AJ455985 AJ455985 |
| 22 | 109.5 | 3.3 | 767 | 1 | AJ637792 AJ637792 |
| 23 | 109 | 3.3 | 726 | 5 | BU350410 603528489 |
| 24 | 108 | 3.3 | 3643 | 9 | AY406378 Pan trogl |
| 25 | 108 | 3.3 | 3647 | 9 | AY406377 Homo sapi |
| 26 | 107 | 3.2 | 699 | 7 | CN396207 170004243 |
| 27 | 107 | 3.2 | 848 | 6 | CA320486 UI-M-FW0- |
| 28 | 107 | 3.2 | 859 | 4 | BG914150 602812668 |
| 29 | 106.5 | 3.2 | 2478 | 3 | AK043970 Mus muscu |
| 30 | 106.5 | 3.2 | 2947 | 3 | AK004820 Mus muscu |
| C 31 | 106 | 3.2 | 987 | 9 | CNS06N86 |
| 32 | 105.5 | 3.2 | 1103 | 5 | BX344037 BX344037 |
| 33 | 105.5 | 3.2 | 3745 | 3 | BC053034 Mus muscu |
| 34 | 105.5 | 3.2 | 5567 | 3 | BC070426 Mus muscu |
| 35 | 105 | 3.2 | 3647 | 9 | AY406379 Mus muscu |
| C 36 | 104.5 | 3.2 | 782 | 7 | CO001759 EST790094 |
| 37 | 104.5 | 3.2 | 947 | 4 | BG332717 602433433 |
| 38 | 104 | 3.1 | 641 | 1 | AJ638144 AJ638144 |
| 39 | 104 | 3.1 | 675 | 6 | CB723823 UI-M-FY0- |
| 40 | 104 | 3.1 | 707 | 5 | BU702565 UI-M-FC0- |
| 41 | 104 | 3.1 | 876 | 9 | CR160126 Forward 8 |
| 42 | 104 | 3.1 | 1009 | 4 | BM449478 AGENCOURT |
| 43 | 104 | 3.1 | 1398 | 9 | CL981728 OsIFSC046 |
| 44 | 103.5 | 3.1 | 629 | 1 | AJ638196 AJ638196 |
| C 45 | 103.5 | 3.1 | 851 | 8 | BH546557 BOGUM60TF |

ALIGNMENTS

RESULT 1
AZ289042
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

489 bp DNA linear GSS 27-JUL-2000
RPCI-23-59A6.TVB RPCI-23 Mus musculus genomic clone RPCI-23-59A6,
genomic survey sequence.
AZ289042 GI:9530917
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 489)
Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
Akinret,B., Levins,M., McGann,S., Tsagaye,G., Geer,K., Krol,M., de
Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other_GSSs: RPCI-23-59A6.TJB
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tldb/bacends/mouse/bac_end_intro.html
Plate: 59 row: A column: 6